一种基于 GLRGMM 的间歇过程在线监控策略 *

赵小强 [a, b, c]，周文伟 [a, b]

(兰州理工大学 a. 电气工程与信息工程学院；b. 甘肃省工业过程先进控制重点实验室；c. 国家级电气与控制工程实验教学中心，兰州 730050)

摘 要：针对间歇过程的非线性和动态性，提出了全局—局部正则化高斯混合模型 (GLRGMM)算法。首先引入邻域保持嵌入算法提取局部流形结构，通过寻求一种低维投影对非线性过程进行全局结构保持，同时最大限度地保留局部流形特征；然后通过对高斯混合模型引入正则项来在线监控更新高斯模型，获取非线性数据流形结构，解决数据动态性问题；最后集成全局—局部监控指标实现在线监控。通过青霉素发酵过程进行了验证，结果表明所提算法比 DPCA、GLNPE 具有更好的在线监控效果。

关键词：间歇过程；在线监控；动态特性；邻域保持嵌入算法；高斯混合模型

中图分类号：TP277 doi: 10.19734/j.issn.1001-3695.2018.04.0441

Online monitoring strategy for batch process based on GLRGMM

Zhao Xiaoqiang[a, b, c], Zhou Wenwei[a, b]

(*a. College of Electrical & Information Engineering, b. Key Laboratory of Gansu Advanced Control for Industrial Processes, c. National Experimental Teaching Center of Electrical & Control Engineering, Lanzhou University of Technology, Lanzhou* 730050, *China*)

Abstract: Aiming at nonlinear and dynamic characteristics of batch process, this paper proposed a global-local regularization Gaussian mixture model (GLRGMM) algorithm. At first, this algorithm introduced Neighborhood Preserving Embedding algorithm to extract the local manifold structure. It preserved global structure of nonlinear process by seeking a low-dimensional projection while preserved local manifold features to the maximum extent at the same time. Then it introduced the regularization term into GMM to online monitor and update Gaussian model. It obtained the manifold structure of nonlinear data and solved the data dynamic characteristic problem simultaneously. Finally, it used the integration of global-local monitoring indicators to effectively achieve online monitoring. The results verified by the penicillin fermentation process show that the proposed algorithm has better online monitoring effect than DPCA and GLNPE algorithms.

Key words: batch process; online monitoring; dynamic characteristic; NPE; GMM

0 引言

间歇过程由于其操作性强和灵活度高等优点成为一种重要的工业生产方式，特别是在精细化工、生物制药和食品饮料等小批量、高附加值产品的生产和制备当中得到广泛应用[1,2]。强非线性、动态性是间歇过程的固有特征，是实际生产过程和在线监控中面临的问题，而寻求一种有效的解决途径对降低经济损失和避免人员伤亡具有重要意义[3]。

随着工业过程控制系统的发展，大量过程数据被获取并得到有效利用，其中以 PCA、ICA、PLS 为主要代表的基于多元统计过程监控(multivariate statistical process monitoring, MSPM)方法在工业生产过程中得到了迅速发展和广泛应用。为了解决间歇过程中的非线性和动态性问题，KPCA[4]、DPCA 和 DPLS[5]应运而生并得到了成功应用。Lee 等人[6]提出了多向核主成分分析算法(multi-way kernel PCA, MKPCA)；Peng 等人[7]在 T-PLS 的基础之上提出了全影核偏最小二乘(total kernel partial leasts quares, T-KPLS)算法。这些算法在一定程度上能有效解决过程变量之间的非线性特性，但仍存在计算量大和核参数选择盲目性问题，并且忽略了时序相关性，没有考虑动态性问题，难以保证过程监控的准确性。Fan 等人[8]提出了基于核函数的动态独立元分析 KDICA 算法；Lu 等人[9]提出了双向动态主成分分析方法。这些方法都是通过历史数据扩展当前数据来消除时序相关性，并不能很好地消除过程数据的动态性。为了同时解决过程数据中的非线性和动态性，Wang 等人[10]提出了双向动态

收稿日期：2018-04-27；**修回日期：**2018-06-08 **基金项目：**国家自然科学基金资助项目（61763029）

作者简介：赵小强（1969-），男，陕西岐山人，教授，博士，博导，主要研究方向为故障诊断、生产调度（xqzhao@lut.cn）；周文伟（1991-），男，硕士研究生，主要研究方向为故障诊断.

核主成分分析法(dynamic kernel PCA, DKPCA)；Zhang 等人[11]提出了递归核 PCA 方法；Jia 等人[12]提出了基于批动态核 PCA 方法。这些方法虽然能够同时监测过程数据非线性和动态特性，但没有考虑过程数据的全局和局部结构信息，仍存在主元个数多、故障误报率较高等问题。

近年来，流形算法由于其优越的局部流形结构特征提取和高维数据处理能力而越来越多地应用于过程监控领域。Hu 等人应用局部保持投影算法(locality preserving projection, LPP)[13]和邻域保持嵌入算法(neighborhood preserving embedding, NPE)[14]对间歇过程进行了有效监控；Zhang 等人[15]提出了一种基于全局—局部结构分析(global-local structure analysis, GLSA)算法，保持了全局结构信息的同时获取了局部流形结构；Yu[16]根据 GLSA 提出了局部—全局 PCA (local and global PCA, LGPCA)算法； Tong 等人[17]提出了多流形投影(multi-manifold projection, MMP)算法，结合了全局和局部方差信息用于过程监控。这些算法虽然提取了数据的全局和局部信息，但并没有有效解决数据的非线性和动态性问题。由于高斯混合模型(Gaussian mixture model, GMM)可以更好地处理复杂过程数据而被用于估计复杂数据分布，并成功应用于非线性、动态特性的工业过程中[18~20]。

为了解决间歇过程的非线性和动态特性，并且有效提取过程数据的全局—局部信息，本文提出了全局—局部正则化高斯混合模型(global-local regularization Gaussian mixture model, GLRGMM)算法。由于 NPE 可以更好地获取高维数据的局部流形特征，通过对 NPE 进行改进求取一个新的低维投影，获取高维空间全局-局部流形结构，实现非线性过程数据降维和特征提取；然后对 GMM 进行动态流形正则化，自动更新获取当前数据；最后通过 NLLP (negative log likelihood probability)和 MD (mahalanobis distance)统计全局、局部概率生成综合监控指标 GLP (global local probability)，实现对间歇过程的在线监控。

1　全局—局部邻域保持嵌入算法

1.1　邻域保持嵌入算法(NPE)

NPE 算法能够挖掘出高维空间中隐藏的低维流形结构，能保持原始空间的局部流形结构。其目的是将 R^D 空间的高维数据映射到一个相对低维的特征空间 $Y=\{y_1,y_2,\cdots,y_d\}$，$\{y_i|y_i\in R^d\}(d<D)$，同时求出相应的变换矩阵 $A=\{a_1,a_2\cdots,a_d\}$。

首先根据 k 近邻法确定样本点 x_i 的 k 个近邻点并组成局部邻域 $M(x_i)$，再通过近邻点来计算权重系数矩阵 W，权重系数矩阵可通过求解式（1）的最优解得到。

$$\min\sum_i\left\|x_i-\sum_j w_{ij}x_j\right\|,\ \sum_j w_{ij}=1,j=1,2,\cdots,N \tag{1}$$

映射变换矩阵 A 通过式（2）最优解得到。

$$A=\arg\min\left[\sum_i\left\|y_i-\sum_j w_{ij}y_j\right\|^2\right]=\arg\min tr\left(A^{\mathrm{T}}XMX^{T}A\right) \tag{2}$$

NPE 虽然能很好地保持局部流形结构，并能揭示隐藏在高维数据中的非线性，但忽略了全局结构，并不能更好地保持数据结构信息。因此，本文首先提出了全局-局部邻域保持嵌入(global-local neighborhood preserving embedding,GLNPE) 算法。

1.2　全局—局部邻域保持嵌入(GLNPE)算法

1.2.1 全局结构保持

对于 R^D 空间样本集 $X=[x_1,x_2,\cdots,x_n]$，寻找投影矩阵 A，通过投影映射 $Y=A^{\mathrm{T}}X$ 将其投影到低维特征空间，根据方差最大方向来最大程度地保留原始数据全局结构信息，其目标函数如式（3）所示。

$$J_{global}(A)=\frac{1}{n}\sum_{i=1}^{n}A^{\mathrm{T}}(x_i-\bar{x})(x_i-\bar{x})^{\mathrm{T}}A=A^{\mathrm{T}}CA \tag{3}$$

其中：$\bar{x}=\frac{1}{n}\sum x_i$；$C=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})^{\mathrm{T}}$。

1.2.2 局部结构保持

局部结构保持是通过每个数据点及其近邻点进行重构来挖掘出数据局部结构信息，其保留局部结构的目标函数如式（4）所示。

$$\begin{aligned}J_{local}(A)&=\min\sum\left\|y_i-\sum_{j=1}^{k}w_{ij}y_i\right\|^2=\min Y^{\mathrm{T}}(I-W)^{\mathrm{T}}(I-W)Y\\&=\min A^{\mathrm{T}}X(I-W)^{\mathrm{T}}(I-W)X^{\mathrm{T}}A=\min A^{\mathrm{T}}XMX^{\mathrm{T}}A\end{aligned} \tag{4}$$

其中：W 为权重矩阵；$M=(I-W)^{\mathrm{T}}(I-W)$。

1.2.3 全局—局部邻域保持嵌入算法(GLNPE)

为了更加充分地提取过程数据特征，需要找到一种投影使得在低维空间同时保留全局和局部信息，即在式（3）最大的同时能保证式（4）最小。为了解决这个复杂优化问题，本文提出了新的目标函数：

$$\begin{aligned}J(A)&=\min\frac{J_{local}(A)}{\eta J_{global}(A)+(1-\eta)J_{local}(A)}\\&=\min\frac{A^{\mathrm{T}}XMX^{\mathrm{T}}A}{\eta A^{\mathrm{T}}CA+(1-\eta)A^{\mathrm{T}}XMX^{\mathrm{T}}A}\end{aligned} \tag{5}$$

其中：$0<\eta<1$ 为可调平衡参数，本文取 $\eta=0.85$。式（5）最小化问题可以通过拉普拉斯方法求得投影矩阵 A：

$$(\eta C+(1-\eta)M)A=\lambda A^{\mathrm{T}}XMX^{\mathrm{T}}A \tag{6}$$

由此可得低维特征空间数据：

$$Y = A^{\mathrm{T}} X \tag{7}$$

通过投影映射 A 将高维数据投影到低维特征空间的同时最大程度地保留了全局—局部结构信息，更有助于对非线性数据特征的提取。

2　全局—局部正则化高斯混合模型(GLRGMM)算法

假设原始变量空间数据为 $D=\{X_i | i=1,2,\cdots,n\}$，其中 n 为变量个数，每个变量的维数为 m，则样本的高斯混合模型可表示为

$$p(X)=\sum_{k=1}^{K}\omega_k g(X|\mu_k,\Sigma_k) \tag{8}$$

其中：$K,\omega_k,\mu_k,\Sigma_k$ 分别表示高斯成分数，第 k 个高斯成分的权重、均值和方差，其目标函数可以写成

$$J(\theta)=\max[\log p(X)]=\max\left[\log\prod_{i=1}^{n}p(x_i|\theta)\right] \tag{9}$$

通过 GLNPE 提取特征空间中的两个数据点 X_i 和 X_j，其余弦相似性矩阵 $S_{ij}=\cos\angle(X_i,X_j)$。

定义正则项 R_k：

$$\begin{aligned}R_k&=\frac{1}{2}\sum_{i,j=1}^{n}\left(p(C_k|X_i)-p(C_k|X_j)\right)^2 S_{ij}\\&=\sum_{i=1}^{n}p(C_k|X_i)^2 I_{ii}-\sum_{i,j=1}^{n}p(C_k|X_i)p(C_k|X_j)S_{ij}\\&=f_{C_k}^{\mathrm{T}}I_{ii}f_{C_k}-f_{C_k}^{\mathrm{T}}Sf_{C_k}\\&=f_{C_k}^{\mathrm{T}}Lf_{C_k}\end{aligned} \tag{10}$$

其中：I_{ii} 为单位矩阵；$L=I-S$；$p(C_k|X_i)$ 为 X_i 属于第 k 个高斯成分的后验概率，且

$$f_k=\left[p(C_k|X_1),\cdots,p(C_k|X_m)\right]^{\mathrm{T}} \tag{11}$$

通过最小化 R_k 确定后验概率，使得不同高斯成分之间平滑过渡，去掉后验概率最小的高斯成分，更新当前高斯成分，实现动态过程平滑过渡和数据实时更新。引入一个包含正则项 R_k 的新目标函数：

$$J_{new}(\theta)=\sum_{i=1}^{n}\sum_{k=1}^{K}p(C_k|X_i)\log\left(\omega_k g(X_i|\mu_k,\Sigma_k)\right)-\lambda R_k \tag{12}$$

其中：λ 为正则化参数。采用 K-means 算法得到初始参数 $K_0,\omega_0,\mu_0,\Sigma_0$ 通过 EM 迭代自动更新确定高斯成分参数 $\theta(K,\omega,\mu,\Sigma)$。

E-step：

通过最小化 R_k 确定后验概率：

$$p^{(s)}(C_k|X_i)=p^{(s-1)}(C_k|X_i)-\frac{\sum_{j=1}^{n}S_{ij}p(C_k|X_j)}{\sum_{j=1}^{n}S_{ij}} \tag{13}$$

M-step:

$$\omega_k^{(s+1)}=\frac{\sum_{i=1}^{n}p^{(s)}(C_k|X_i)}{n} \tag{14}$$

$$\mu_k^{(s+1)}=\frac{\sum_{i=1}^{n}x_i p^{(s)}(C_k|X_i)}{\sum_{i=1}^{n}p^{(s)}(C_k|X_i)} \tag{15}$$

$$\Sigma_k^{(s+1)}=\frac{\sum_{i=1}^{n}p^{(s)}(C_k|X_i)\left(X_i-\mu_k^{(s+1)}\right)\left(X_i-\mu_k^{(s+1)}\right)^{\mathrm{T}}}{\sum_{i=1}^{n}p^{(s)}(C_k|X_i)} \tag{16}$$

其中：$\omega_k^{(s+1)},\mu_k^{(s+1)},\Sigma_k^{(s+1)}$ 分别表示第 $s+1$ 步迭代中第 k 个高斯成分的权重、均值和方差。

GLRGMM 通过 GLNPE 对非线性数据的全局—局部结构特征进行了更加充分地提取，更符合实际间歇过程数据特性。其详细的算法流程如图 1 所示。

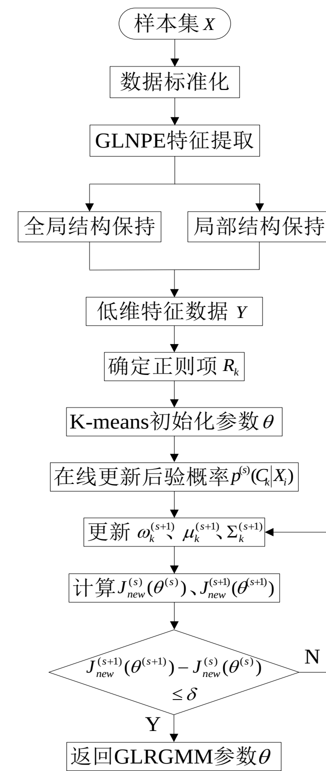


图 1　GLRGMM 算法流程

Fig.1　Flow chart of GLRGMM algorithm

3　基于 GLRGMM 的监控策略

3.1　间歇过程数据预处理

间歇过程实时监测得到的数据为三维数据：

$X(I\times J\times K)$，其中：I 为批次个数；J 为变量个数；K 为采样时间。在实际的生产过程中，由于不同批次的采样时间不同等原因而不可能做到完全重复同批次的生产，这就不可避免地出现数据不等长问题。本文采用最短长度法[21]，以最短的一个批次数据作为标准而将其余批次数据进行硬性截取，使得所有批次数据都具有相等长度。

对间歇过程的三维数据进行建模，需要把已处理成相等长度的数据展开成与续过程相类似的二维数据。本文结合了批次展开可以去除非线性、变量展开可以消除数据预估问题的优点，展开方式如图 2 所示。

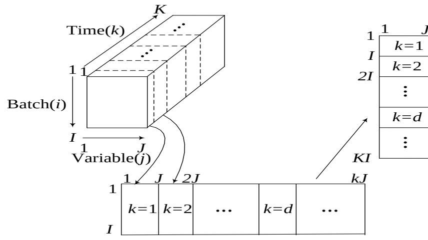


图 2　间歇过程数据展开方式

Fig.2　Scheme of batch process data unfolding

首先，将三维数据 $X(I\times J\times K)$ 按批次方向展开成二维数据 $X(I\times JK)$ 并进行标准化处理：

$$\bar{x}_{j,k}=\frac{1}{I}\sum_{i=1}^{I}x_{i,j,k} \tag{17}$$

$$\delta_{j,k}=\sqrt{\frac{1}{I-1}\sum_{i=1}^{I}(x_{i,j,k}-\bar{x}_{j,k})^2} \tag{18}$$

$$\bar{x}_{i,j,k}=\frac{x_{i,j,k}-\bar{x}_{j,k}}{\delta_{j,k}} \tag{19}$$

得到的数据在一定程度上可以消除非线性和动态特性；然后将标准化的二维数据 $X(I\times JK)$ 重置为三维数据 $X(I\times J\times K)$，再按变量方向展开成二维数据 $X(KI\times J)$ 并进行标准化。这种展开方式可以消除在线监控时的数据预估问题，提高了实时监控的精确率，降低了误报率。

3.2　GLP(global local probability)监控指标

为了统计全局概率指标，首先引入负对数似然概率（negative log likelihood probability, NLLP）。

$$NLLP(X_i)=-\log p(C_k|X_i) \tag{20}$$

全局概率指标通过比较测试数据与训练数据的 NLLP 得到

$$P_g(X_{test})=\Pr\{NLLP(X_{test})\le NLLP(X_{train})\} \tag{21}$$

引入服从 χ^2 分布的马氏距离（Mahalanobis distance, MD）：

$$MD(X_i)=(X_i-\mu_k)(\Sigma_k)^{-1}(X_i-\mu_k)^{\mathrm{T}} \tag{22}$$

定义局部概率指标：

$$\begin{aligned}P_l^{(k)}(X_{test})=\Pr\{&MD((X_{train},C_k)|X_{train}\in C_k)\\&\le MD((X_{test},C_k)|X_{test}\in C_k)\}\end{aligned} \tag{23}$$

通过集成全局—局部概率生成 GLP(global local probability)监控指标：

$$GLP=\sum_{k=1}^{K}P_g(C_k|X_{test})P_l^{(k)}(X_{test}) \tag{24}$$

3.3　离线建模

a）在正常工况下采集若干批次数据 $X(I\times J\times K)$ 作为训练样本，采用 3.1 节的方法进行数据预处理，最终得到已标准化处理的二维数据 $X(KI\times J)$；

b）通过 GLNPE 进行全局—局部信息保持，由式（6）（7）得到低维特征空间投影数据 Y；

c）利用 GLRGMM 对数据 Y 进行正则化，用 EM 算法确定参数 θ；

d）通过式（20）（22）分别计算 $NLLP(X_{train})$、$MD(X_{train})$。

3.4　在线监控

a）将测试数据按 3.1 节的方法进行数据预处理。

b）通过 GLNPE 算法而得到新的测试数据的投影数据 Y。

c）通过 GLRGMM 对新的投影数据 Y 进行正则化，用 EM 算法确定参数 θ。

d）通过式（20）（22）分别计算 $NLLP(X_{test})$、$MD(X_{test})$。

e）通过式（21）（23）分别计算全局、局部概率指标。

f）通过式（24）计算 GLP 监控指标，若 GLP>0.99，则出现故障；否则，重复步骤 b）~f）直到出现故障报警或生产过程结束。

4　案例研究

青霉素作为一种重要抗生素而被临床广泛应用，其生产制备过程是典型的间歇过程，具有强非线性、动态性等特性。本文基于美国 Illinois 州立理工学院开发的 Pensim 2.0 标准青霉素仿真平台[22]产生出间歇过程数据，它可产生出不同初始条件和不同工况下青霉素发酵过程中各变量每个时刻的数据用于分析研究。

本文将每批次的反应时间设定为 400 h，采样时间设为 0.5 h，在正常范围设置不同初始条件，不引入故障的情况下共产生 35 个批次正常工况下数据，从产生的 18 个变量数据中选择其

中 10 个过程变量作为监控变量（表 1），构成三维数据 X(35×10×800)作为训练样本。

表 1　过程变量

Table 1　Process variables

变量号	变量名称
1	通风速率（L/h）
2	搅拌功率（w）
3	补料温度（K）
4	溶解氧浓度（mmol/L）
5	反应器体积（L）
6	排气二氧化碳浓度（mmol/L）
7	pH 值
8	发酵罐浓度（K）
9	产生热（K）
10	冷水流加速率（L/h）

Pensim 2.0 提供了三种故障类型（通风速率、搅拌功率、底物流加速率）。本文引入故障类型 1（即变量 1 通风速率）和故障类型 2（即变量 2 搅拌功率），在时间 200~300 h(即采样点 400~600)加入 15%的阶跃信号作为故障信号，所产生的数据作为故障测试数据以供在线监控。通过 GLRGMM 对训练数据建模得到高斯成分如图 3 所示。分别用 DPCA、GLNPE 和 GLRGMM 算法进行监控，对于 DPCA 和 GLNPE 给出故障 1 的监控图，如图 4~7 所示；其余详细监控结果见表 2、3。GLRGMM 对故障 1 和 2 的监控结果如图 8、9 所示。

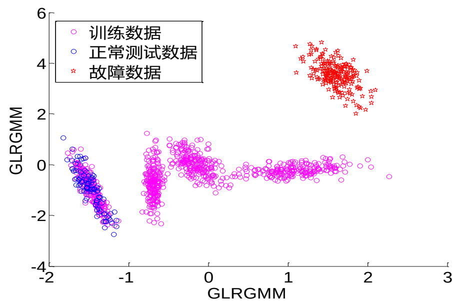


图 3　训练数据建模结果

Fig.3　Modeling results of training data

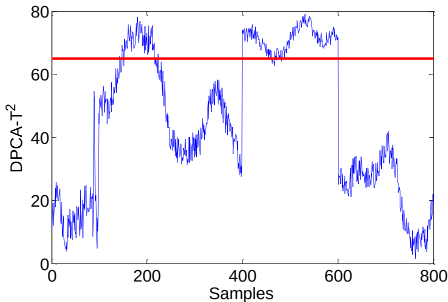


图 4　DPCA 对故障 1 的 T^2 监控图

Fig.4　T^2 monitoring diagram of DPCA for fault 1

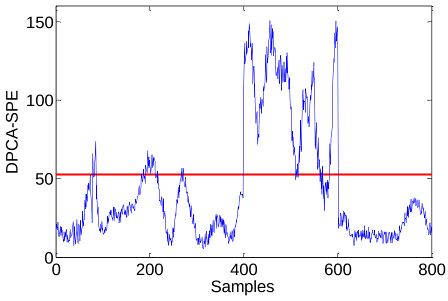


图 5　DPCA 对故障 1 的 SPE 监控图

Fig.5　SPE monitoring diagram of DPCA for fault 1

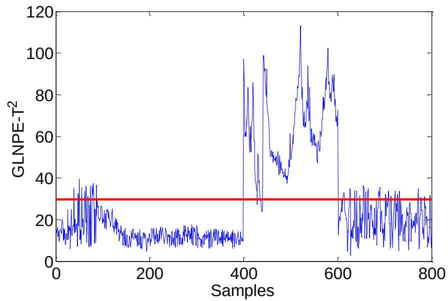


图 6　GLNPE 对故障 1 的 T^2 监控图

Fig.6　T^2 monitoring diagram of GLNPE for fault 1

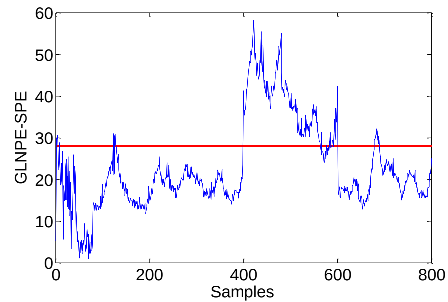


图 7　GLNPE 对故障 1 的 SPE 监控图

Fig.7　SPE monitoring diagram of GLNPE for fault 1

表 2　不同方法对故障 1 的监控统计　　/%

Table 2　Monitoring statistics of different methods for fault 1　　/%

方法（统计量）	DPCA（T^2）	DPCA（SPE）	GLNPE（T^2）	GLNPE（SPE）	GLRGMM（GLP）
误报率	26.36	10.7	3.02	9.18	0.5
漏报率	19.4	16.92	3.98	11.44	0
检测率	80.6	83.08	96.02	88.56	100

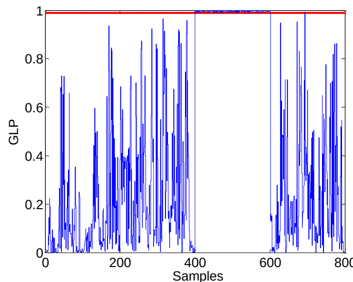


图 8　GLRGMM 对故障 1 的监控图

Fig.8　Monitoring diagram of GLRGMM for fault 1

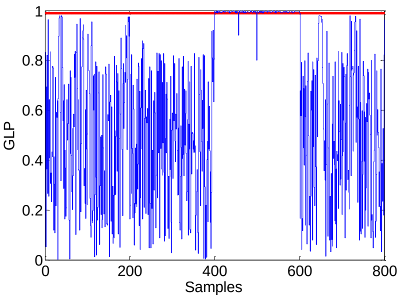


图 9　GLRGMM 对故障 2 的监控图

Fig.9　Monitoring diagram of GLRGMM for fault 2

表 3　不同方法对故障 2 的监控统计　/%

Table 3　Monitoring statistics of different methods for fault 2　/%

方法（统计量）	DPCA（T^2）	DPCA（SPE）	GLNPE（T^2）	GLNPE（SPE）	GLRGMM（GLP）
误报率	16.48	8.16	14.85	6.36	0
漏报率	18.91	15.53	14.43	12.97	8.96
检测率	81.09	84.47	85.57	87.03	91.04

图 4~ 7 为采用 DPCA 和 GLNPE 对故障 1 的监控，从图中可以明显看出 DPCA 和 GLNPE 均有误报、漏报产生。虽然 DPCA 在一定程度上可以解决动态性；但从表 2 可以看出 GLNPE 的误报率、漏报率较低、检测率高于 DPCA，这是因为 DPCA 并不能很好地保持非线性数据过程中的全局、局部结构，只是对线性过程的动态特性处理。GLNPE 不仅保持了全局结构信息，而且可以有效地提取隐藏在高维空间中的低维流形结构，虽然不是对非线性最有效地解决途径，但通过保持非线性数据的全局、局部信息，故障检测率明显优于传统直接忽略非线性的方法。表 3 为故障 2 的监控统计结果，从表 3 也可以看出 GLNPE 比 DPCA 具有更好的监控效果，进一步证实了保持全局-局部信息对非线性过程监控的有效性。

图 8、9 为 GLRGMM 对故障 1 和故障 2 的监控，可以看出 GLRGMM 的监控效果优于 DPCA 和 GLNPE。GLRGMM 可以检测出所有故障 1 的变量，没有漏报产生，误报率仅为 0.5%，检测速度快；对故障 2 虽有漏报，但没有误报，检测率为 91.04%，故障敏感性高。这是因为 GLRGMM 不仅最大程度地保持了非线性数据的全局、局部结构，而且考虑了数据的动态性，对非线性和动态特性过程具有明显的优越性。

5　结束语

强非线性、动态性是间歇过程的固有特性，本文提出的基于 GLRGMM 的监控策略，不仅通过改进 NPE 对非线性数据进行全局—局部结构保持，而且通过引入正则项对高斯模型进行在线更新，最后整合全局—局部概率指标进行在线监控。通过青霉素发酵过程的仿真实验结果表明该方法比 DPCA、GLNPE 有更好的监控效果，降低了误报率、漏报率，提高了故障监控的快速性和精确性。

参考文献：

[1] 郭金玉，袁堂明，李元. 一种不等长的多模态间歇过程故障检测方法 [J]. 化工学报, 2016, 67 (7): 2916-2924. (Guo Jinyu, Yuan Tangming, Li Yuan. Fault detection method for uneven-length multimode batch processes [J]. CIESC Journal, 2016, 67 (7): 2916-2924.)

[2] 王普，李春蕾，高学金，等. 基于角结构统计量的 MKECA 间歇过程故障监测 [J]. 仪器仪表学报, 2017, 38 (1): 174-180. (Wang Pu, Li Chunlei, Gao Xuejin, *et al.* Research on batch process monitoring method based on multi-way kernel entropy component analysis and angular structure statistic [J]. Chinese Journal of Scientific Instrument, 2017, 38 (1): 174-180.)

[3] 赵小强，王涛. 基于 TGNPE 算法的间歇过程故障诊断 [J]. 化工学报, 2016, 67 (3): 1055-1062. (Zhao Xiaoqiang, Wang Tao. Batch process fault diagnosis based on TGNPE algorithm [J]. CIESC Journal, 2016, 67 (3): 1055-1062.)

[4] Ben I K, Liama M, Weihs C. Variable window adaptive kernel principal component analysis for nonlinear nonstationary process monitoring [J]. Computers & Industrial Engineering, 2011, 61 (3): 437-466.

[5] Chen J, Liu K C. On-line batch process monitoring using dynamic PCA and dynamic PLS models [J]. Chemical Engineering Science, 2002, 57 (1): 63-75.

[6] Lee J M, Changkyoo Y, In-Beum L. Fault detection of batch processes using multi-way kernel principal component analysis [J]. Computers & Chemical Engineering, 2004, 28 (9): 1837-1847.

[7] Peng Kaixiang, Zhang Kai, Li Gang, *et al.* Contribution rate plot for nonlinear quality related fault diagnosis with application to the hot strip mill process [J]. Control Engineering Practice, 2013, 21 (4): 360-369.

[8] Fan Jicong, Wang Youqing. Fault detection and diagnosis of non-linear non-Gaussian dynamic processes using kernel dynamic independent component analysis [J]. Information Sciences, 2014, 259 (3): 369-379.

[9] Lu Ningyun, Yao Yuan, Gao Furong, *et al.* Two-dimensional dynamic PCA for batch process monitoring [J]. AIChE Journal: Chemical Engineering Research and Design, 2005, 51 (12): 3300-3304.

[10] Wang Ting, Wang Xiaogang, Zhang Yingwei, *et al.* Fault detection of nonlinear dynamic processes using dynamic kernel principal component analysis [C]// Proc of the 7th World Congress on Intelligent Control and Automation. 2008: 3009-3014.

[11] Zhang Yingwei, Li Shuai, Teng Yongdong. Dynamic processes monitoring using recursive kernel principal component analysis [J]. Chemical Engineering Science, 2012, 72 (16): 78-86.

[12] Jia Mingxing, Chu Fei, Wang Fuli, *et al.* On-line batch process monitoring using batch dynamic kernel principal component analysis [J]. Chemometrics and Intelligent Laboratory Systems, 2010, 101 (2): 110-122.

[13] Hu Kunlun, Yuan Jingqi. Multivariate statistical process control based on multiway locality preserving projections [J]. Journal of Process Control,

2008, 18 (7-8): 797-807.

[14] Hu Kunlun, Yuan Jingqi. Statistical monitoring of fed-batch process using dynamic multiway neighborhood preserving embedding [J]. Chemometrics & Intelligent Laboratory Systems. 2008, 90 (2): 195-203.

[15] Zhang Muguang, Ge Zhiqiang, Song Zhihuan, *et al*. Global-local structure analysis model and its application for fault detection and identification [J]. Industrial & Engineering Chemistry Research, 2011, 50 (11): 6837-6848.

[16] Yu Jianbo. Local and global principal component analysis for process monitoring [J]. Journal of Process Control, 2012, 22 (7): 1358-1373.

[17] Tong Chudong, Yan Xuefeng. Statistical process monitoring based on a multi-manifold projection algorithm [J]. Chemometrics & Intelligent Laboratory Systems, 2014, 130 (15): 20–28.

[18] Yu Jie, Qin S J. Multimode process monitoring with Bayesian inference-based finite Gaussian mixture models [J]. AIChE Journal, 2008, 54 (7): 1811-1829.

[19] Chen Tao, Zhang Jie. On-line multivariate statistical monitoring of batch processes using Gaussian mixture model [J]. Computers & Chemical Engineering, 2010, 34 (4): 500-507.

[20] Yu Jianbo. Semiconductor manufacturing process monitoring using Gaussian mixture model and Bayesian method with local and nonlocal information [J]. IEEE Trans on Semiconduct, Manufact, 2012, 25 (3): 480–493.

[21] Kourti T. Multivariate dynamic data modeling for analysis and statistical process control of batch processes, startups and grade transitions [J]. Journal of Chemometrics, 2003, 17 (1): 93-109.

[22] Birol G, Ündey C, Cinar A. A modular simulation package for fed-batch fermentation: penicillin production [J]. Computers and Chemical Engineering, 2002, 26 (11): 1553-1565.